

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

(i) APPLICANT: Coleman, Roger  
Stuart, Susan G.

(ii) TITLE OF THE INVENTION: A NOVEL HUMAN JAK2 KINASE

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
- (B) STREET: 3174 Porter Drive
- (C) CITY: Palo Alto
- (D) STATE: CA
- (E) COUNTRY: US
- (F) ZIP: 94304

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: DOS
- (D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/567,508
- (B) FILING DATE: 05-DEC-1995
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Billings, Lucy J.
- (B) REGISTRATION NUMBER: 36,749
- (C) REFERENCE/DOCKET NUMBER: PF-0049US

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 650-855-0555
- (B) TELEFAX: 650-845-4166

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4482 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Placenta
- (B) CLONE: 179527

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCACGCGTCC	GGTTGCCAAC	CCGCAGGCGA	CTGGGCGCTT	CATCCCACCC	TCACCCCTT	60
CCAGCCAAGG	TGGCTGATCG	GAGTCAGGCT	CTCGAGGTGCG	CATTGCCACG	AAACGGNGTG	120
TGTGAGCGCG	TTGTCcccGG	NCCCCGGGGC	CACTTCCCT	CGGCCTAGNA	GACTGGACTG	180
GGGAAGGACG	GGTCTGTTGT	ACCCGGGAGG	TGGAAGGAAA	AGCCGAAAGC	GGAGAAGTGT	240
GCGGGAGGGG	AGTCTCCGCG	CGGAGGNAGA	CCGGNCTCCT	CCAGTGCAGG	TTGTGCGCTG	300
GGGAGCCAGC	CASGGCAAAT	GTTCTGAAAAA	AGACTCTGCA	TGGGAATGGC	CTGCCTTACG	360
ATGACAGAAA	TGGAGGGAAC	ATCCACCTCT	TCTATATATC	AGAATGGTGA	TATTTCTGGA	420
AATGCCAATT	CTATGAAGCA	AATAGATCCA	GTTCTTCAGG	TGTATCTTA	CCATTCCCTT	480
GGGAAATCTG	AGGCAGATT	TCTGACCTT	CCATCTGGG	AGTATGTTGG	AGAAGAAATC	540
TGTATTGCTG	CTTCTAAAGC	TTGTGGTATC	ACACCTGTG	ATCATAATAT	GTTTGCTTTA	600
ATGAGTGAAGA	CAGAAAGGAT	CTGGTATCCA	CCCAACCATG	TCTTCCATAT	AGATGAGTCA	660
ACCAGGCATA	ATGTACTCTA	CAGAAATAAGA	TTTACTTTC	CTCGTTGGTA	TTGCAGTGGC	720
AGCAACAGAG	CCTATCGGCA	TGGAATATCT	CGAGGTGCTG	AAGCTCCTCT	TCTTGATGAC	780
TTTGTATGT	CTTACCTCTT	TGCTCAGTGG	CGGCATGATT	TTGTGCATGG	ATGGATAAAA	840
GTACCTGTGA	CTCATGAAAC	ACAGGAAGAA	TGTCTTGGG	TGACAGTGT	AGATATGATG	900
AGAATAGCCA	AAGAAAACGA	TCAAAACCCCA	CTGGCCATCT	ATAACTCTAT	CAGCTACAAG	960
ACATTCTAC	CAAAATGTAT	TCGAGCAAAG	ATCCAAGACT	ATCATATTT	GACAAGGAAG	1020
CGAATAAGGT	ACAGATTTCG	CAGATTATT	CAGCAATTCA	GCCATGCAA	AGCCACTGCC	1080
AGAAACTTGA	AACTTAAGTA	TCTTATAAAT	CTGGAAACTC	TGCACTCTGC	CTTCTACACA	1140
GAGAAATTG	AAGTAAAAGA	ACCTGGAAGT	GGTCCTTCAG	GTGAGGAGAT	TTTGCAACC	1200
ATTATAATAA	CTGGAAACGG	TGGAATTCA	TGGTCAAGAG	GGAAACATAA	AGAAAGTGAG	1260
ACACTGACAG	AACAGGATT	ACAGTTATAT	TGCGATTTTC	CTAATATTAT	TGATGTCAGT	1320
ATTAAGCAAG	CAAACCAAGA	GGGTTCAAAT	GAAAGCCGAG	TTGTAACAT	CCATAAGCAA	1380
GATGGTAAAAA	ATCTGGAAAT	TGAACCTTAGC	TCATTAAGGG	AAGCTTGTG	TTTCGTGTCA	1440
TTAATTGATG	GATATTATAG	ATTAACGTCA	GATGCACATC	ATTACCTCTG	TAAAGAAGTA	1500
GCACCTCCAG	CCGTGCTTGA	AAATATACAA	AGCAACTGTC	ATGGCCCAAT	TTCGATGGAT	1560
TTTGCCTTA	GTAAACTGAA	GAAAGCAGGT	AATCAGACTG	GACTGTATGT	ACTTCGATGC	1620
AGTCTTAAGG	ACTTTAATAA	ATATTTTG	ACTTTGTG	TCGAGCGAGA	AAATGTCTT	1680
GAATATAAAC	ACTGTTGAT	TACAAAAAAAT	GAGAATGAAG	AGTACAACCT	CAGTGGGACA	1740
AAGAAGAATC	TCAGCAGTCT	TAAGATCTT	TGAAATTGTT	ACCAGATGGA	AACTGTTCGC	1800
TCAGACAATA	TAATTTCCA	GTAACTTAA	TGCTGTCCCC	CAAAGCCAAA	AGATAAAATCA	1860
AACCTCTAG	TCTTCAGAAC	GAATGGTGT	TCTGATGTAC	CAACCTCACC	AACATTACAG	1920
AGGCCTACTC	ATATGAACCA	AATGGTGT	TACAAAATCA	GAAATGAAGA	TTTGATATT	1980
AATGAAAGCC	TTGCCAAGG	CACTTTACA	AAAGATTTTA	AAGGCGTAGC	AAGAGAAGTA	2040
GGAGACTACG	GTCAACTGCA	TGAAACAGAA	GTTCTTTAA	AAAGTCTGGA	TAAGCACAC	2100
AGGAACATT	CAGACTCTT	CTTGAAAGCA	GCAAGTATGA	TGAGCAAGCT	TTCTCACAA	2160
CATTGGTTT	TAATTATGG	AGTATGTGTC	TGTGGAGACG	AGAATATTCT	GGTCAGGAG	2220
TTTGTAAAAT	TTGATCACT	AGATACATAT	CTGAAAAGA	ATAAAAATTG	TATAATATA	2280
TTATGGAAAC	TTGAAGTTGC	TAACAGTT	GCATGGGCCA	TGCATTTCT	AGAAGAAAAC	2340
ACCCTTATTC	ATGGAATGT	ATGTGCCAA	AAATATTCTGC	TTATCAGAGA	AGAAGACAGG	2400
AAGACAGGAA	ATCCTCTTT	CATCAAACCT	AGTGATCCTG	GCATTAGTAT	TACAGTTTG	2460
CCAAAGGACA	TTCTTCAGGA	GAGAATACCA	TGGGTACAC	CTGAATGCAT	TGAAATCTCT	2520
AAAAATTAA	ATTGGCACAC	AGACAAATGG	AGTTTGGTA	CCACTTTGTG	GGAAATCTGC	2580
AGTGGAGGAG	ATAAACCTCT	AAGTGTCTG	GATTCTCAA	GAAAGCTACA	ATTTTATGAA	2640
GATAGGCATC	AGCTTCTGC	ACCAAAGTGG	GCAGAATTG	CAAACCTTAT	AAATAATTGT	2700
ATGGATTATG	AACCAAGATT	CAGGCTCT	TTCAAGGCCA	TCATACGAGA	TCTTAACAGT	2760
TTGTTTACTC	CAGATTATGA	ACTTAAACA	GAAAATGACA	TGTTACAAA	TATGAGGATA	2820
GGTGCCTTGG	GGTTTCTGG	TGCCCTTGA	GACCGGGATC	CTACACAGTT	TGAAGAGAGA	2880
CATTGAAAT	TTCTACAGCA	ACTTGGCAAG	GGTAATTTC	GGAGTGTGGA	GATGTGCCGG	2940
TATGACCTC	TACAGGACAA	CACTGGGAG	GTGGTCGCTG	AAAAAAAGCT	TCAGCATAGT	3000
ACTGAAGAGC	ACCTAAGAGA	CTTGAAAGG	GAAATTGAAA	TCCTGAAATC	CCTACAGCAT	3060
GACAACATTG	TAAGTACAA	GGGAGTGTG	TACAGTGTG	GTCGGCGTAA	TCTAAAATTA	3120
ATTATGGAAT	ATTTACCAT	TGGAAGTTA	CGAGACTATC	TTCAAAACA	TAAAGAACGG	3180
ATAGATCACA	TAAAACCTCT	GCAGTACACA	TCTCAGATAT	GCAAGGGTAT	GGAGTATCTT	3240
GGTACAAAAA	GGTATATCCA	CAGGGATCTG	GCAACGAGAA	ATATATTGGT	GGAGAACGGAG	3300
ACAGAGTTA	AAATTGGRGA	TTTGGTTA	ACCAAAGTCT	TGCCACAAGA	CAAAGAATAC	3360
TATAAAGTAA	AAGAACCTGG	TGAAAGTCCC	ATATTCTGGT	ATGCTCCAGA	ATCACTGACA	3420
GAGAGCAAGT	TTTCTGTGGC	CTCAGATGTT	TGGAGCTTTG	GAGTGGTTCT	GTATGAACTT	3480
TTCACATACA	TTGAGAAGAG	TAAAAGTCCA	CCAGCGGAAT	TTATGCGTAT	GATTGGCAAT	3540
GACAAACAAG	GACAGATGAT	CGTGTCCAT	TTGATAGAAC	TTTGAAAGAA	TAATGGAAGA	3600
TTACCAAGAC	CAGATGGAT	CCCAGATGAG	ATCTATATGA	TCATGACAGA	ATGCTGGAAC	3660
ATAATGTAA	ATCAACGCC	CTCCTTCTAGG	GATCTAGCTC	TTCGAGTGG	TCAAATAAGG	3720
GATAACATGG	CTGGATGAAA	GAAATGACCT	TCATTCTGAG	ACCAAAGTAG	ATTTACAGAA	3780
CAAAGTTTA	TATTTCACAT	TGCTGTGGAC	TATTATTACA	TATATCATT	TTATATAAAT	3840
CATGATGCTA	GCCAGCAAAG	ATGTGAAAAT	ATCTGCTCAA	AACTTCAAA	GTTCAGTAAG	3900

TTTTTCTTCA	TGAGGCCACC	AGTAAAAGAC	ATTAATGAGA	ATTCCCTTAGC	AAGGATTGG	3960
TAAGAAGTTT	CTTAAACATT	GTCAGTTAAC	ATCACTCTG	TCTGGCAAAA	GAAAAAAAAT	4020
AGACTTTTC	AACTCAGCTT	TTTGAGACCT	GAAARAATTA	TTATGTAAAT	TTTGCATGT	4080
TAAAGATGCA	CAGAATATGT	ATGTATAGTT	TTTACCCACAG	TGGATGTATA	ATACCTTGGC	4140
ATCTTGTGTG	ATGTTAACCA	CACATGAGGG	CTGGTGTTC	TTAATACTGT	TTTCTAATT	4200
TTCCATGGTT	AATCTATAAT	TAATTACTTC	ACTAAACAAA	CAAATTAAGA	TGTTCAGATA	4260
ATTGAATAAG	TACCTTTGTG	TCCTTGTCA	TTTATATCGC	TGGCCAGCAT	TATAAGCAGG	4320
TGTATACTT	TAGCTTGTAG	TTCCATGTAC	TGTAAATATT	TTTCACATAA	AGGAAACAAA	4380
TGTCTAGTT	TATTTGTATA	GGAAATTGTC	CCTGACCCCTA	AATAATACAT	TTTGAAATGA	4440
AACAAGCTTA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AG		4482

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1132 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY:
- (B) CLONE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Gly	Met	Ala	Cys	Leu	Thr	Met	Thr	Glu	Met	Gly	Thr	Ser	Thr	
1				5			10			15					
Ser	Ser	Ile	Tyr	Gln	Asn	Gly	Asp	Ile	Ser	Gly	Asn	Ala	Asn	Ser	Met
			20				25				30				
Lys	Gln	Ile	Asp	Pro	Val	Leu	Gln	Val	Tyr	Leu	Tyr	His	Ser	Leu	Gly
	35				40						45				
Lys	Ser	Glu	Ala	Asp	Tyr	Leu	Thr	Phe	Pro	Ser	Gly	Glu	Tyr	Val	Gly
	50				55					60					
Glu	Glu	Ile	Cys	Ile	Ala	Ala	Ser	Lys	Ala	Cys	Gly	Ile	Thr	Pro	Val
	65				70				75			80			
Tyr	His	Asn	Met	Phe	Ala	Leu	Met	Ser	Glu	Thr	Glu	Arg	Ile	Trp	Tyr
			85					90				95			
Pro	Pro	Asn	His	Val	Phe	His	Ile	Asp	Glu	Ser	Thr	Arg	His	Asn	Val
			100				105					110			
Leu	Tyr	Arg	Ile	Arg	Phe	Tyr	Phe	Pro	Arg	Trp	Tyr	Cys	Ser	Gly	Ser
	115					120					125				
Asn	Arg	Ala	Tyr	Arg	His	Gly	Ile	Ser	Arg	Gly	Ala	Glu	Ala	Pro	Leu
	130					135				140					
Leu	Asp	Asp	Phe	Val	Met	Ser	Tyr	Leu	Phe	Ala	Gln	Trp	Arg	His	Asp
	145				150				155			160			
Phe	Val	His	Gly	Trp	Ile	Lys	Val	Pro	Val	Thr	His	Glu	Thr	Gln	Glu
			165					170				175			
Glu	Cys	Leu	Gly	Met	Thr	Val	Leu	Asp	Met	Met	Arg	Ile	Ala	Lys	Glu
	180					185					190				
Asn	Asp	Gln	Thr	Pro	Leu	Ala	Ile	Tyr	Asn	Ser	Ile	Ser	Tyr	Lys	Thr
		195					200					205			
Phe	Leu	Pro	Gln	Cys	Ile	Arg	Ala	Lys	Ile	Gln	Asp	Tyr	His	Ile	Leu
	210					215					220				
Thr	Arg	Lys	Arg	Ile	Arg	Tyr	Arg	Phe	Arg	Arg	Phe	Ile	Gln	Gln	Phe
	225					230				235			240		
Ser	Gln	Cys	Lys	Ala	Thr	Ala	Arg	Asn	Leu	Lys	Leu	Lys	Tyr	Leu	Ile
			245				250					255			
Asn	Leu	Glu	Thr	Leu	Gln	Ser	Ala	Phe	Tyr	Thr	Glu	Lys	Phe	Glu	Val
		260					265					270			
Lys	Glu	Pro	Gly	Ser	Gly	Pro	Ser	Gly	Glu	Glu	Ile	Phe	Ala	Thr	Ile
	275					280					285				
Ile	Ile	Thr	Gly	Asn	Gly	Gly	Ile	Gln	Trp	Ser	Arg	Gly	Lys	His	Lys
	290					295					300				
Glu	Ser	Glu	Thr	Leu	Thr	Glu	Gln	Asp	Leu	Gln	Leu	Tyr	Cys	Asp	Phe
	305					310				315			320		
Pro	Asn	Ile	Ile	Asp	Val	Ser	Ile	Lys	Gln	Ala	Asn	Gln	Glu	Gly	Ser
							325			330			335		
Asn	Glu	Ser	Arg	Val	Val	Thr	Ile	His	Lys	Gln	Asp	Gly	Lys	Asn	Leu
					340			345				350			
Glu	Ile	Glu	Leu	Ser	Ser	Leu	Arg	Glu	Ala	Leu	Ser	Phe	Val	Ser	Leu
			355				360				365				
Ile	Asp	Gly	Tyr	Tyr	Arg	Leu	Thr	Ala	Asp	Ala	His	His	Tyr	Leu	Cys
	370					375					380				

Lys Glu Val Ala Pro Pro Ala Val Leu Glu Asn Ile Gln Ser Asn Cys  
 385 390 395 400  
 His Gly Pro Ile Ser Met Asp Phe Ala Ile Ser Lys Leu Lys Lys Ala  
 405 410 415  
 Gly Asn Gln Thr Gly Leu Tyr Val Leu Arg Cys Ser Pro Lys Asp Phe  
 420 425 430  
 Asn Lys Tyr Phe Leu Thr Phe Ala Val Glu Arg Glu Asn Val Ile Glu  
 435 440 445  
 Tyr Lys His Cys Leu Ile Thr Lys Asn Glu Asn Glu Glu Tyr Asn Leu  
 450 455 460  
 Ser Gly Thr Lys Lys Asn Phe Ser Ser Leu Lys Asp Leu Leu Asn Cys  
 465 470 475 480  
 Tyr Gln Met Glu Thr Val Arg Ser Asp Asn Ile Ile Phe Gln Phe Thr  
 485 490 495  
 Lys Cys Cys Pro Pro Lys Pro Lys Asp Lys Ser Asn Leu Leu Val Phe  
 500 505 510  
 Arg Thr Asn Gly Val Ser Asp Val Pro Thr Ser Pro Thr Leu Gln Arg  
 515 520 525  
 Pro Thr His Met Asn Gln Met Val Phe His Lys Ile Arg Asn Glu Asp  
 530 535 540  
 Leu Ile Phe Asn Glu Ser Leu Gly Gln Gly Thr Phe Thr Lys Ile Phe  
 545 550 555 560  
 Lys Gly Val Arg Arg Glu Val Gly Asp Tyr Gly Gln Leu His Glu Thr  
 565 570 575  
 Glu Val Leu Leu Lys Val Leu Asp Lys Ala His Arg Asn Tyr Ser Glu  
 580 585 590  
 Ser Phe Phe Glu Ala Ala Ser Met Met Ser Lys Leu Ser His Lys His  
 595 600 605  
 Leu Val Leu Asn Tyr Gly Val Cys Val Cys Gly Asp Glu Asn Ile Leu  
 610 615 620  
 Val Gln Glu Phe Val Lys Phe Gly Ser Leu Asp Thr Tyr Leu Lys Lys  
 625 630 635 640  
 Asn Lys Asn Cys Ile Asn Ile Leu Trp Lys Leu Glu Val Ala Lys Gln  
 645 650 655  
 Leu Ala Trp Ala Met His Phe Leu Glu Glu Asn Thr Leu Ile His Gly  
 660 665 670  
 Asn Val Cys Ala Lys Asn Ile Leu Leu Ile Arg Glu Glu Asp Arg Lys  
 675 680 685  
 Thr Gly Asn Pro Pro Phe Ile Lys Leu Ser Asp Pro Gly Ile Ser Ile  
 690 695 700  
 Thr Val Leu Pro Lys Asp Ile Leu Gln Glu Arg Ile Pro Trp Val Pro  
 705 710 715 720  
 Pro Glu Cys Ile Glu Asn Pro Lys Asn Leu Asn Leu Ala Thr Asp Lys  
 725 730 735  
 Trp Ser Phe Gly Thr Thr Leu Trp Glu Ile Cys Ser Gly Gly Asp Lys  
 740 745 750  
 Pro Leu Ser Ala Leu Asp Ser Gln Arg Lys Leu Gln Phe Tyr Glu Asp  
 755 760 765  
 Arg His Gln Leu Pro Ala Pro Lys Trp Ala Glu Leu Ala Asn Leu Ile  
 770 775 780  
 Asn Asn Cys Met Asp Tyr Glu Pro Asp Phe Arg Pro Ser Phe Arg Ala  
 785 790 795 800  
 Ile Ile Arg Asp Leu Asn Ser Leu Phe Thr Pro Asp Tyr Glu Leu Leu  
 805 810 815  
 Thr Glu Asn Asp Met Leu Pro Asn Met Arg Ile Gly Ala Leu Gly Phe  
 820 825 830  
 Ser Gly Ala Phe Glu Asp Arg Asp Pro Thr Gln Phe Glu Glu Arg His  
 835 840 845  
 Leu Lys Phe Leu Gln Gln Leu Gly Lys Gly Asn Phe Gly Ser Val Glu  
 850 855 860  
 Met Cys Arg Tyr Asp Pro Leu Gln Asp Asn Thr Gly Glu Val Val Ala  
 865 870 875 880  
 Val Lys Lys Leu Gln His Ser Thr Glu Glu His Leu Arg Asp Phe Glu  
 885 890 895  
 Arg Glu Ile Glu Ile Leu Lys Ser Leu Gln His Asp Asn Ile Val Lys  
 900 905 910  
 Tyr Lys Gly Val Cys Tyr Ser Ala Gly Arg Arg Asn Leu Lys Leu Ile

915		920		925											
Met	Glu	Tyr	Leu	Pro	Tyr	Gly	Ser	Leu	Arg	Asp	Tyr	Leu	Gln	Lys	His
930						935						940			
Lys	Glu	Arg	Ile	Asp	His	Ile	Lys	Leu	Leu	Gln	Tyr	Thr	Ser	Gln	Ile
945						950					955				960
Cys	Lys	Gly	Met	Glu	Tyr	Leu	Gly	Thr	Lys	Arg	Tyr	Ile	His	Arg	Asp
						965			970				975		
Leu	Ala	Thr	Arg	Asn	Ile	Leu	Val	Glu	Asn	Glu	Asn	Arg	Val	Lys	Ile
						980			985				990		
Gly	Asp	Phe	Gly	Leu	Thr	Lys	Val	Leu	Pro	Gln	Asp	Lys	Glu	Tyr	Tyr
						995		1000				1005			
Lys	Val	Lys	Glu	Pro	Gly	Glu	Ser	Pro	Ile	Phe	Trp	Tyr	Ala	Pro	Glu
						1010		1015			1020				
Ser	Leu	Thr	Glu	Ser	Lys	Phe	Ser	Val	Ala	Ser	Asp	Val	Trp	Ser	Phe
025						1030					1035				1040
Gly	Val	Val	Leu	Tyr	Glu	Leu	Phe	Thr	Tyr	Ile	Glu	Lys	Ser	Lys	Ser
						1045			1050				1055		
Pro	Pro	Ala	Glu	Phe	Met	Arg	Met	Ile	Gly	Asn	Asp	Lys	Gln	Gly	Gln
						1060			1065				1070		
Met	Ile	Val	Phe	His	Leu	Ile	Glu	Leu	Leu	Lys	Asn	Asn	Gly	Arg	Leu
						1075		1080				1085			
Pro	Arg	Pro	Asp	Gly	Cys	Pro	Asp	Glu	Ile	Tyr	Met	Ile	Met	Thr	Glu
						1090		1095				1100			
Cys	Trp	Asn	Asn	Asn	Val	Asn	Gln	Arg	Pro	Ser	Phe	Arg	Asp	Leu	Ala
105						1110					1115				1120
Leu	Arg	Val	Asp	Gln	Ile	Arg	Asp	Asn	Met	Ala	Gly				
						1125					1130				

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1129 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Gly Met Ala Cys Leu Thr Met Thr Glu Met Glu Ala Thr Ser Thr  
 1 5 10 15  
 Ser Pro Val His Gln Asn Gly Asp Ile Pro Gly Ser Ala Asn Ser Val  
 20 25 30  
 Lys Gln Ile Glu Pro Val Leu Gln Val Tyr Leu Tyr His Ser Leu Gly  
 35 40 45  
 Gln Ala Glu Gly Glu Tyr Leu Lys Phe Pro Ser Gly Glu Tyr Val Ala  
 50 55 60  
 Glu Glu Ile Cys Val Ala Ala Ser Lys Ala Cys Gly Ile Thr Pro Val  
 65 70 75 80  
 Tyr His Asn Met Phe Ala Leu Met Ser Glu Thr Glu Arg Ile Trp Tyr  
 85 90 95  
 Pro Pro Asn His Val Phe His Ile Asp Glu Ser Thr Arg His Asp Ile  
 100 105 110  
 Leu Tyr Arg Ile Arg Phe Tyr Phe Pro His Trp Tyr Cys Ser Gly Ser  
 115 120 125  
 Ser Arg Thr Tyr Arg Tyr Gly Val Ser Arg Gly Ala Glu Ala Pro Leu  
 130 135 140  
 Leu Asp Asp Phe Val Met Ser Tyr Leu Phe Val Gln Trp Arg His Asp  
 145 150 155 160  
 Phe Val His Gly Trp Ile Lys Val Pro Val Thr His Glu Thr Gln Glu  
 165 170 175  
 Glu Cys Leu Gly Met Ala Val Leu Asp Met Met Arg Ile Ala Lys Glu  
 180 185 190  
 Lys Asp Gln Thr Pro Leu Ala Val Tyr Asn Ser Val Ser Tyr Lys Thr  
 195 200 205  
 Phe Leu Pro Lys Cys Val Arg Ala Lys Ile Gln Asp Tyr His Ile Leu  
 210 215 220  
 Thr Arg Lys Arg Ile Arg Tyr Arg Phe Arg Arg Phe Ile Gln Gln Phe  
 225 230 235 240  
 Ser Gln Cys Lys Ala Thr Ala Arg Asn Leu Lys Leu Lys Tyr Leu Ile  
 245 250 255  
 Asn Leu Glu Thr Leu Gln Ser Ala Phe Tyr Thr Glu Gln Phe Glu Val  
 260 265 270  
 Lys Glu Ser Ala Arg Gly Pro Ser Gly Glu Glu Ile Phe Ala Thr Ile  
 275 280 285  
 Ile Ile Thr Gly Asn Gly Gly Ile Gln Trp Ser Arg Gly Lys His Lys  
 290 295 300  
 Glu Ser Glu Thr Leu Thr Glu Gln Asp Val Gln Leu Tyr Cys Asp Phe  
 305 310 315 320  
 Pro Asp Ile Ile Asp Val Ser Ile Lys Gln Ala Asn Gln Glu Cys Ser  
 325 330 335  
 Asn Glu Ser Arg Ile Val Thr Val His Lys Gln Asp Gly Lys Val Leu  
 340 345 350  
 Glu Ile Glu Leu Ser Ser Leu Lys Glu Ala Leu Ser Phe Val Ser Leu  
 355 360 365  
 Ile Asp Gly Tyr Tyr Arg Leu Thr Ala Asp Ala His His Tyr Leu Cys  
 370 375 380  
 Lys Glu Val Ala Pro Pro Ala Val Leu Glu Asn Ile His Ser Asn Cys  
 385 390 395 400  
 His Gly Pro Ile Ser Met Asp Phe Ala Ile Ser Lys Leu Lys Lys Ala  
 405 410 415  
 Gly Asn Gln Thr Gly Leu Tyr Val Leu Arg Cys Ser Pro Lys Asp Phe  
 420 425 430  
 Asn Lys Tyr Phe Leu Thr Phe Ala Val Glu Arg Glu Asn Val Ile Glu

435	440	445
Tyr Lys His Cys Leu Ile Thr	Lys Asn Glu Asn Gly Glu Tyr Asn Leu	
450	455	460
Ser Gly Thr Asn Arg Asn Phe Ser Asn Leu	Lys Asp Leu Leu Asn Cys	
465	470	475
Tyr Gln Met Glu Thr Val Arg Ser Asp Ser	Ile Ile Phe Gln Phe Thr	480
485	490	495
Lys Cys Cys Pro Pro Lys Pro Lys Asp Lys Ser Asn Leu Leu Val Phe		
500	505	510
Arg Thr Asn Gly Ile Ser Asp Val Gln Ile Ser Pro Thr Leu Gln Arg		
515	520	525
His Asn Asn Val Asn Gln Met Val Phe His Lys Ile Arg Asn Glu Asp		
530	535	540
Leu Ile Phe Asn Glu Ser Leu Gly Gln Gly Thr Phe Thr Lys Ile Phe		
545	550	555
Lys Gly Val Arg Arg Glu Val Gly Asp Tyr Gly Gln Leu His Lys Thr		560
565	570	575
Glu Val Leu Leu Lys Val Leu Asp Lys Ala His Arg Asn Tyr Ser Glu		
580	585	590
Ser Phe Phe Glu Ala Ala Ser Met Met Ser Gln Leu Ser His Lys His		
595	600	605
Leu Val Leu Asn Tyr Gly Val Cys Val Cys Gly Glu Glu Asn Ile Leu		
610	615	620
Val Gln Glu Phe Val Lys Phe Gly Ser Leu Asp Thr Tyr Leu Lys Lys		
625	630	635
Asn Lys Asn Ser Ile Asn Ile Leu Trp Lys Leu Gly Val Ala Lys Gln		640
645	650	655
Leu Ala Trp Ala Met His Phe Leu Glu Glu Lys Ser Leu Ile His Gly		
660	665	670
Asn Val Cys Ala Lys Asn Ile Leu Leu Ile Arg Glu Glu Asp Arg Arg		
675	680	685
Thr Gly Asn Pro Pro Phe Ile Lys Leu Ser Asp Pro Gly Ile Ser Ile		
690	695	700
Thr Val Leu Pro Lys Asp Ile Leu Gln Glu Arg Ile Pro Trp Val Pro		
705	710	715
Pro Glu Cys Ile Glu Asn Pro Lys Asn Leu Asn Leu Ala Thr Asp Lys		720
725	730	735
Trp Ser Phe Gly Thr Thr Leu Trp Glu Ile Cys Ser Gly Gly Asp Lys		
740	745	750
Pro Leu Ser Ala Leu Asp Ser Gln Arg Lys Leu Gln Phe Tyr Glu Asp		
755	760	765
Lys His Gln Leu Pro Ala Pro Lys Trp Thr Glu Leu Ala Asn Leu Ile		
770	775	780
Asn Asn Cys Met Asp Tyr Glu Pro Asp Phe Arg Pro Ala Phe Arg Ala		
785	790	795
Val Ile Arg Asp Leu Asn Ser Leu Phe Thr Pro Asp Tyr Glu Leu Leu		800
805	810	815
Thr Glu Asn Asp Met Leu Pro Asn Met Arg Ile Gly Ala Leu Gly Phe		
820	825	830
Ser Gly Ala Phe Glu Asp Arg Asp Pro Thr Gln Phe Glu Glu Arg His		
835	840	845
Leu Lys Phe Leu Gln Gln Leu Gly Lys Gly Asn Phe Gly Ser Val Glu		
850	855	860
Met Cys Arg Tyr Asp Pro Leu Gln Asp Asn Thr Gly Glu Val Val Ala		
865	870	880
Val Lys Lys Leu Gln His Ser Thr Glu Glu His Leu Arg Asp Phe Glu		
885	890	895
Arg Glu Ile Glu Ile Leu Lys Ser Leu Gln His Asp Asn Ile Val Lys		
900	905	910
Tyr Lys Gly Val Cys Tyr Ser Ala Gly Arg Arg Asn Leu Arg Leu Ile		
915	920	925
Met Glu Tyr Leu Pro Tyr Gly Ser Leu Arg Asp Tyr Leu Gln Lys His		
930	935	940
Lys Glu Arg Ile Asp His Lys Lys Leu Leu Gln Tyr Thr Ser Gln Ile		
945	950	955
Cys Lys Gly Met Glu Tyr Leu Gly Thr Lys Arg Tyr Ile His Arg Asp		960
965	970	975

Leu Ala Thr Arg Asn Ile Leu Val Glu Asn Glu Asn Arg Val Lys Ile  
 980 985 990  
 Gly Asp Phe Gly Leu Thr Lys Val Leu Pro Gln Asp Lys Glu Tyr Tyr  
 995 1000 1005  
 Lys Val Lys Glu Pro Gly Glu Ser Pro Ile Phe Trp Tyr Ala Pro Gln  
 1010 1015 1020  
 Ser Leu Thr Glu Ser Lys Phe Ser Val Ala Ser Asp Val Trp Ser Phe  
 025 1030 1035 1040  
 Gly Val Val Leu Tyr Glu Leu Phe Thr Tyr Ile Glu Lys Ser Lys Ser  
 1045 1050 1055  
 Pro Pro Val Glu Phe Met Arg Met Ile Gly Asn Asp Lys Gln Gly Gln  
 1060 1065 1070  
 Met Ile Val Phe His Leu Ile Glu Leu Leu Lys Ser Asn Gly Arg Leu  
 1075 1080 1085  
 Pro Arg Pro Glu Gly Cys Pro Asp Glu Ile Tyr Val Ile Met Thr Glu  
 1090 1095 1100  
 Cys Trp Asn Asn Asn Val Ser Gln Arg Pro Ser Phe Arg Asp Leu Ser  
 105 1110 1115 1120  
 Phe Gly Trp Ile Lys Cys Gly Thr Val  
 1125

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGCGGAAGT GCTCTCGCG GAAG

24

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGTGTGCTAC AGTGCTGGTC GTCG

24